暨南大學
JINAN UNIVERSITY

ANN-Presentation

JBJI - Computational Intelligence

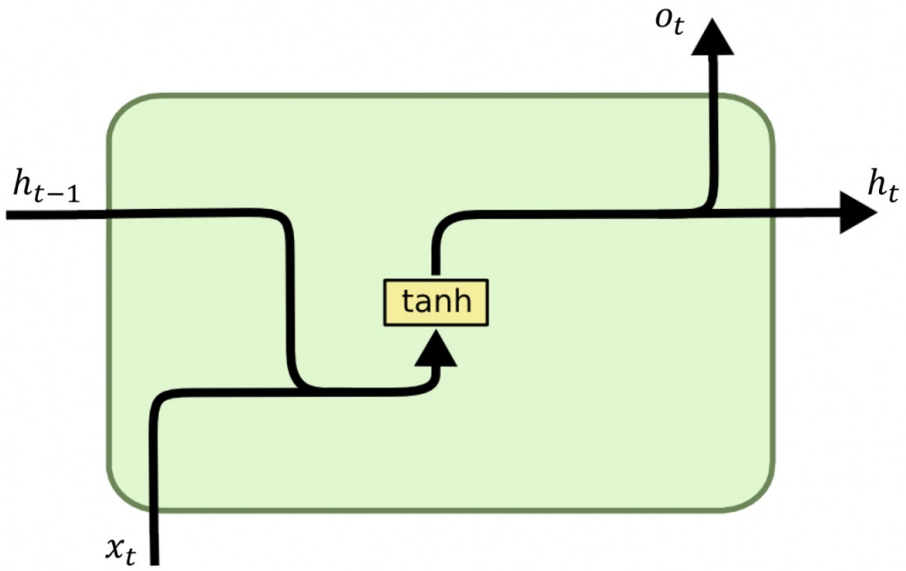
洪梓晋 - 2020101911 - 12/01/2023

CONTENTS

01

RNN-based Model

Recurrent Neural Network



Feed-Forward

$$h_t = \sigma_h(i_t) = \sigma_h(U_h x_t + V_h h_{t-1} + b_h)$$

$$y_t = \sigma_y(a_t) = \sigma_y(W_y h_t + b_h)$$

Backpropagation

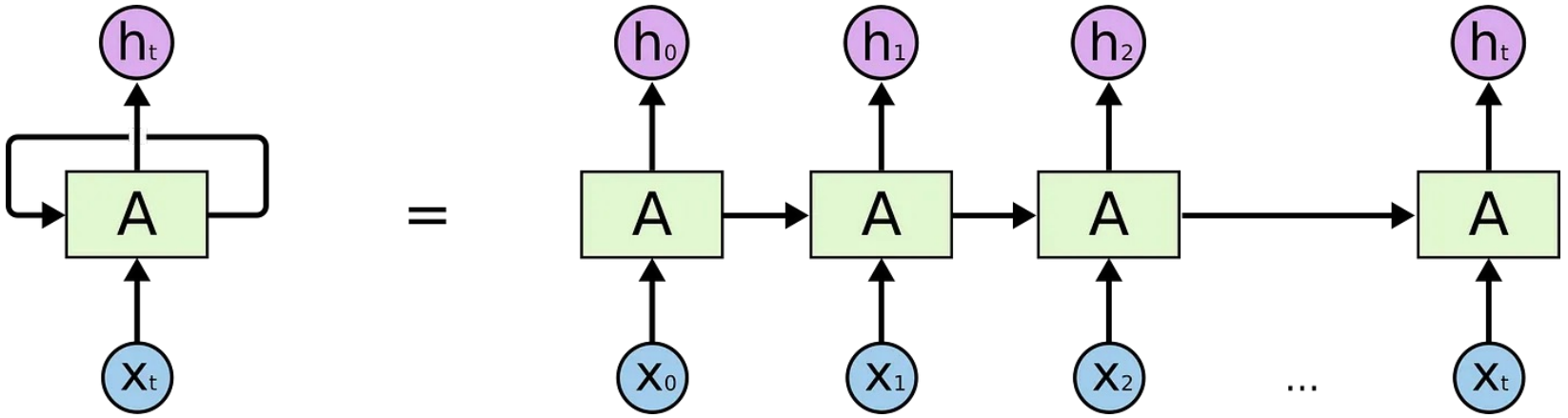
$$\Pi_t = \frac{\partial E_t}{\partial o_t}\frac{\partial o_t}{\partial h_t} + \frac{\partial h_{t+1}}{\partial h_t}\Pi_{t+1}$$

$$\beta_t^U = \beta_{t+1}^U + \Pi_t \frac{\partial h_t}{\partial U_t}$$

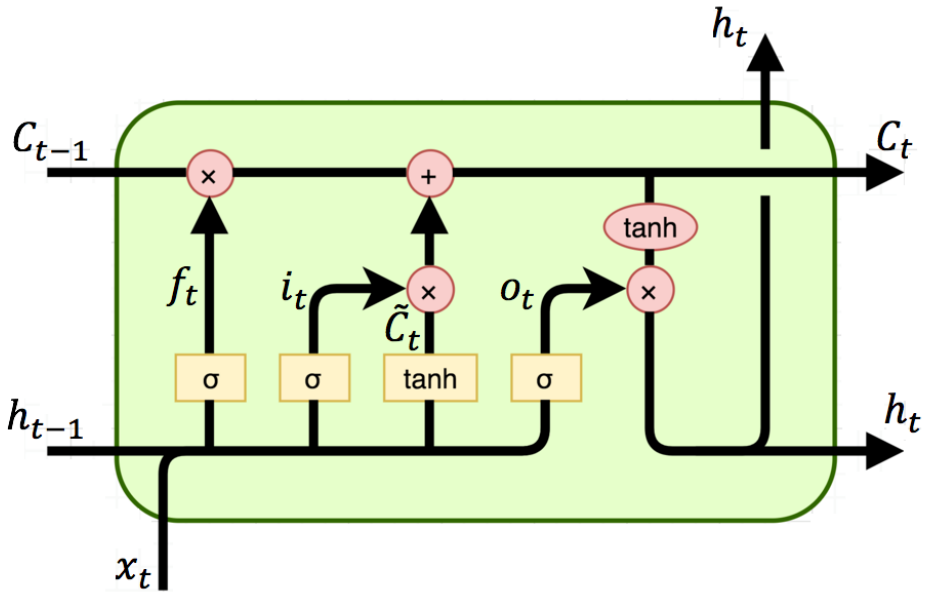
$$\beta_t^V = \beta_{t+1}^V + \Pi_t \frac{\partial h_t}{\partial V_t}$$

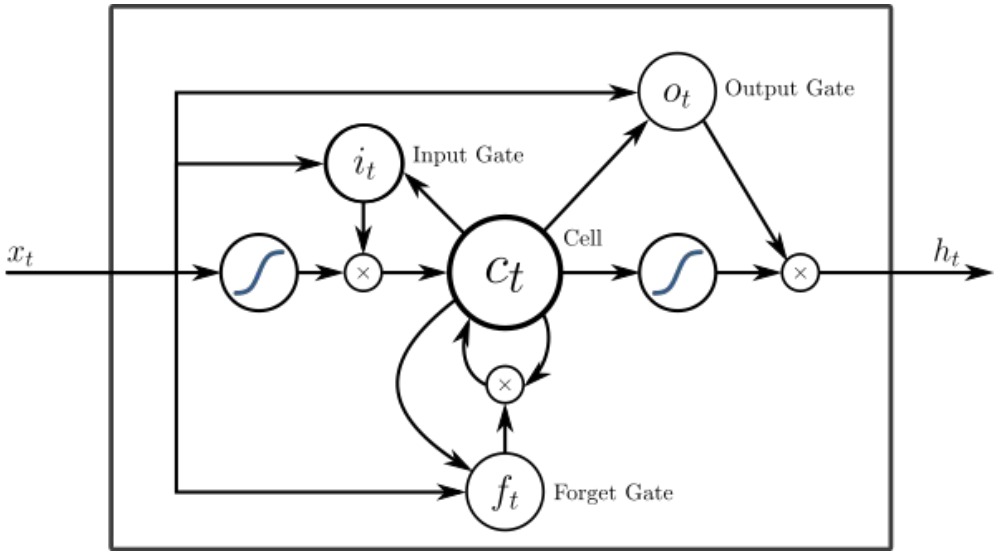
$$\beta_t^W = \beta_{t+1}^W + \frac{\partial E_t}{\partial o_t}\frac{\partial o_t}{\partial W_t}$$

$$\frac{\partial E}{\partial X} \equiv \beta_0^x$$



5 Long-Short Term Memory





Feed-Forward

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

$$h_t = o_t \odot \tanh(C_t)$$

Backpropagation

$$\frac{\partial C_{t+1}}{\partial h_t} = \frac{\partial C_{t+1}}{\partial \tilde{C}_{t+1}} \frac{\partial \tilde{C}_{t+1}}{\partial h_t} + \frac{\partial C_{t+1}}{\partial f_{t+1}} \frac{\partial f_{t+1}}{\partial h_t} + \frac{\partial C_{t+1}}{\partial t_{t+1}} \frac{\partial i_{t+1}}{\partial h_t}$$

$$\frac{\partial C_{t+1}}{\partial C_t}$$

$$\frac{\partial h_{t+1}}{\partial C_t} = \frac{\partial h_{t+1}}{\partial C_{t+1}} \frac{\partial C_{t+1}}{\partial C_t}$$

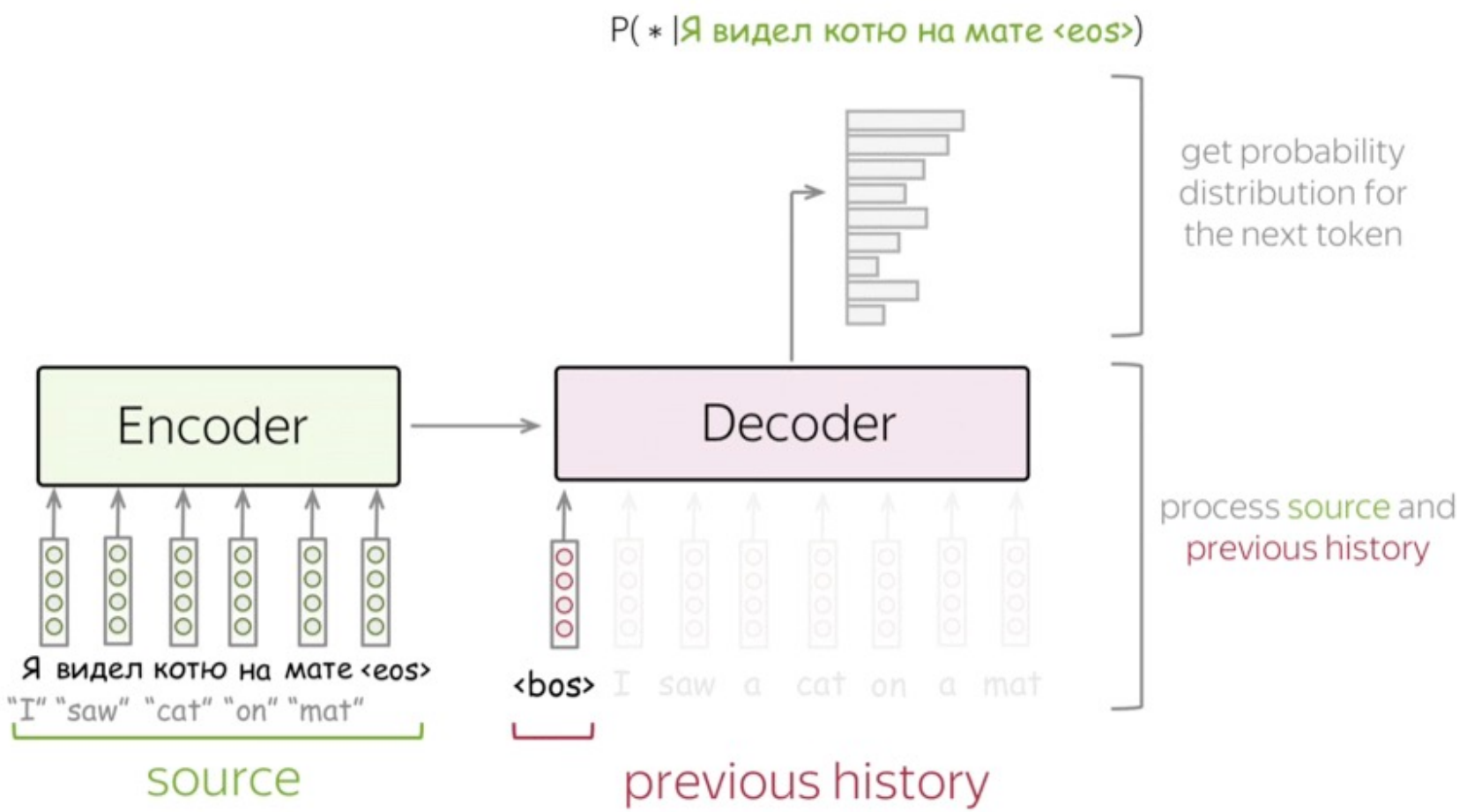
$$\frac{\partial h_{t+1}}{\partial h_t} = \frac{\partial h_{t+1}}{\partial C_{t+1}} \frac{\partial C_{t+1}}{\partial h_t} + \frac{\partial h_{t+1}}{\partial o_{t+1}} \frac{\partial o_{t+1}}{\partial h_t}$$

1. **Input Gate**: Determines which parts of the new information should be added to the cell state. It filters the information through a set of operations, allowing only important information to pass through this gate.
2. **Forget Gate**: Decides which information in the cell state should be discarded or retained. If the forget gate is closed, information is retained; if open, it is forgotten.
3. **Output Gate**: Determines which parts of the cell state should be output to the next hidden layer. It filters the cell state to determine the final output values.

02

Seq2Seq Structure

Sequence to Sequence



Human Translation

$$y^* = \arg\max_y p(y|x)$$

The "probability" is intuitive and is given by a human translator's expertise

Machine Translation

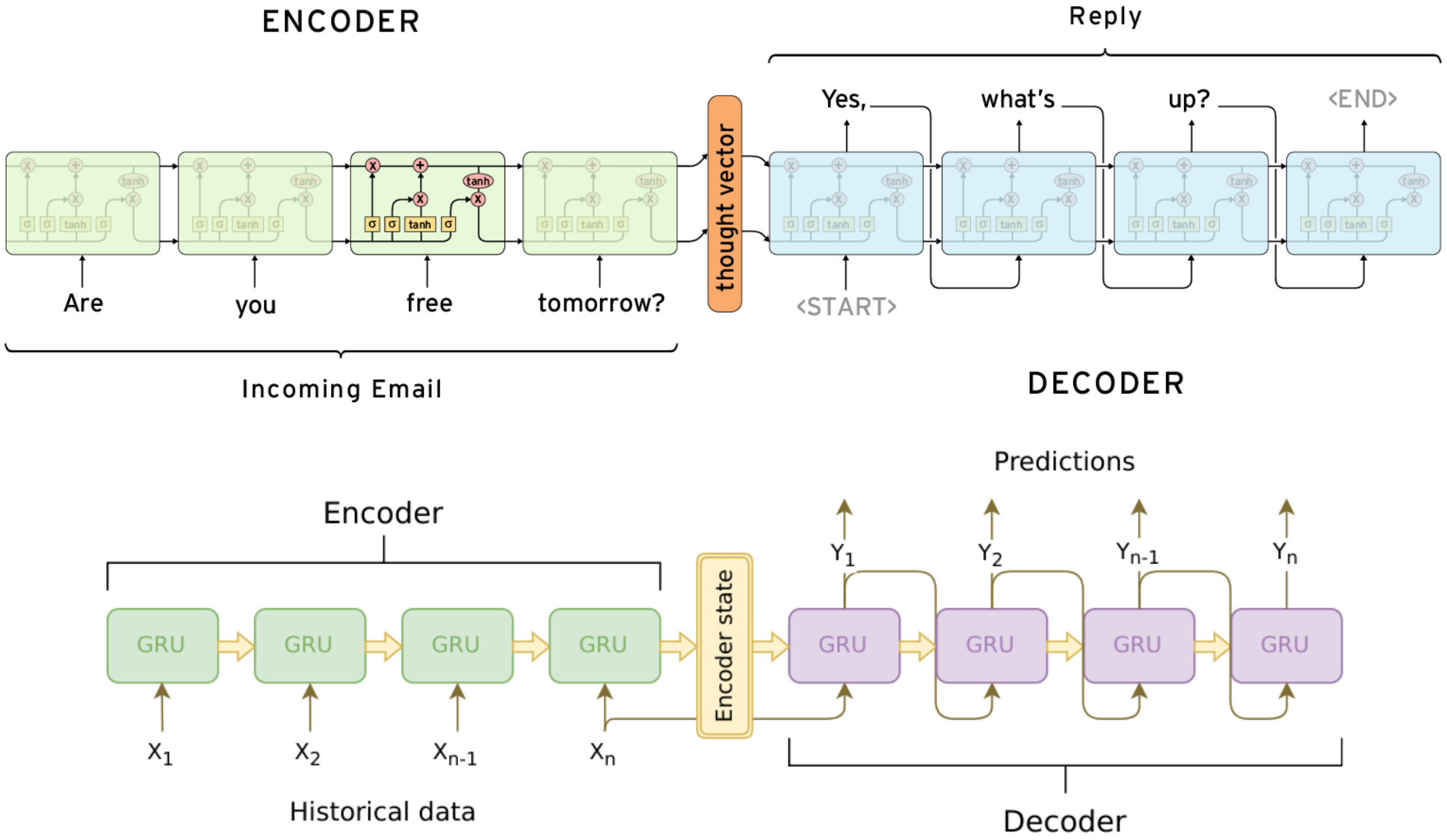
$$y' = \arg\max_y p(y|x,\theta)$$

model parameters

$$Loss(p^*, p) = -p^* \log(p) = -\sum_{i=1}^{|V|} p_i^* \log(p_i).$$

$$Loss(p^*, p) = -\log(p_{y_t}) = -\log(p(y_t|y_{<t}, x)).$$

Sequence to Sequence



Time Series Prediction

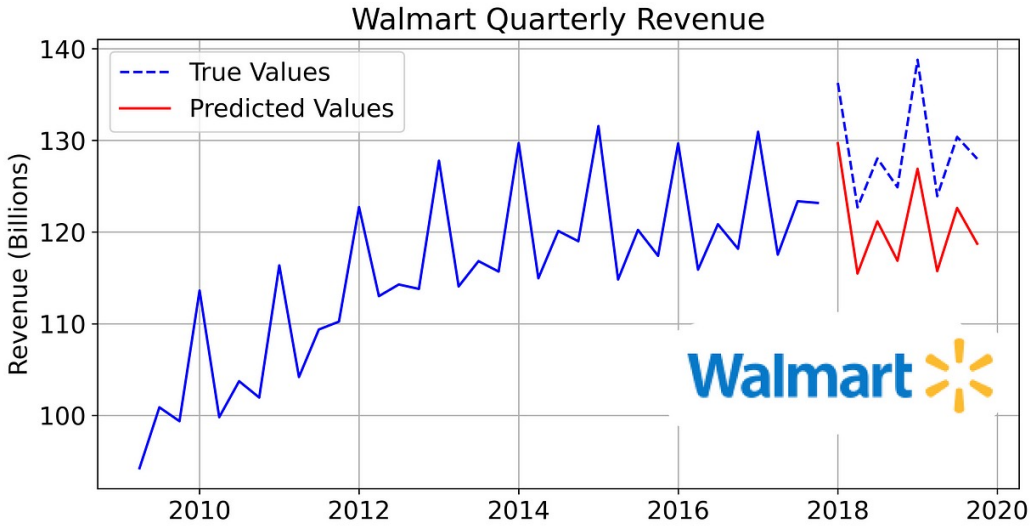
Background

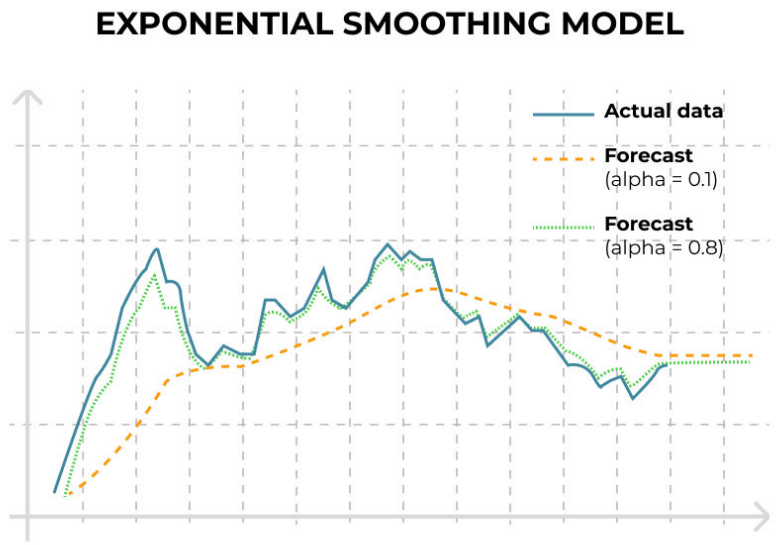
2022 MCM
Problem C: Trading Strategies

Background
Market traders buy and sell volatile assets frequently, with a goal to maximize their total return. There is usually a commission for each purchase and sale. Two such assets are gold and bitcoin.

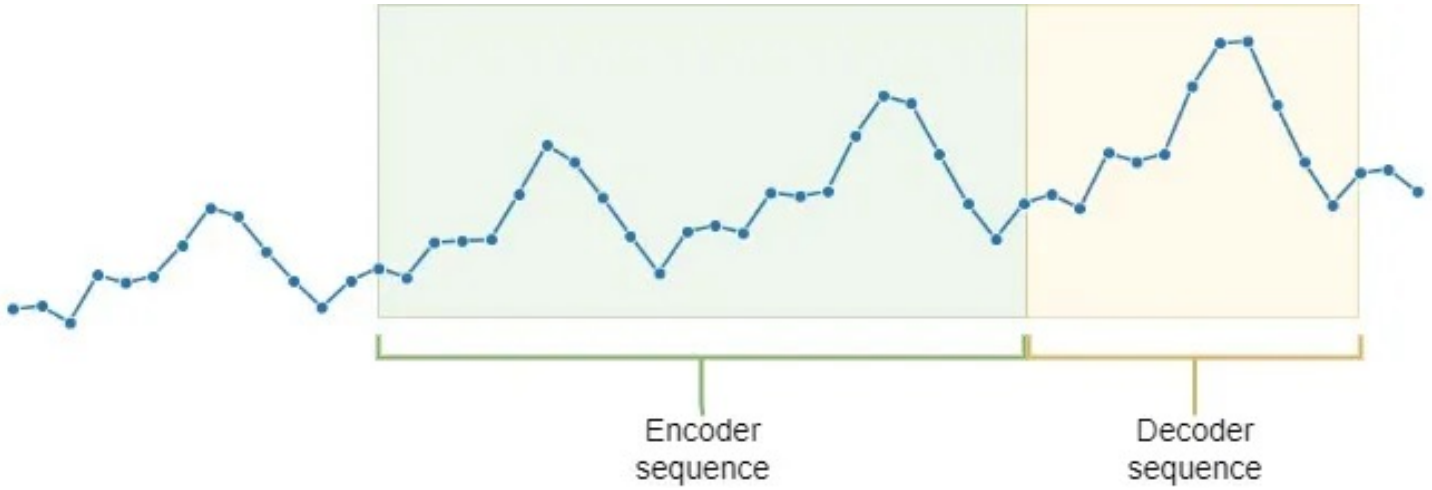
Figure 1: Gold daily prices, U.S. dollars per troy ounce. Source: London Bullion Market Association, 9/11/2021

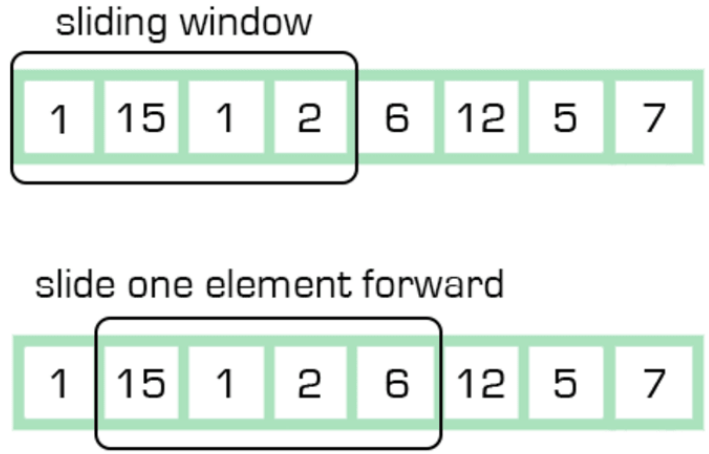
Figure 2: Bitcoin daily prices, U.S. dollars per bitcoin. Source: NASDAQ, 9/11/2021





Sequential Data





Seq2Seq For TS Prediction

Note:

X_t → time series value at timestep **t**

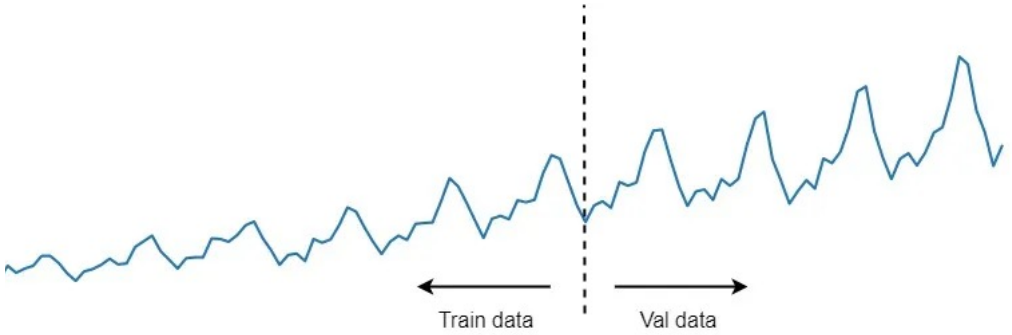
Targets: [X_{t+1} , X_{t+2} , X_{t+3} , ...]

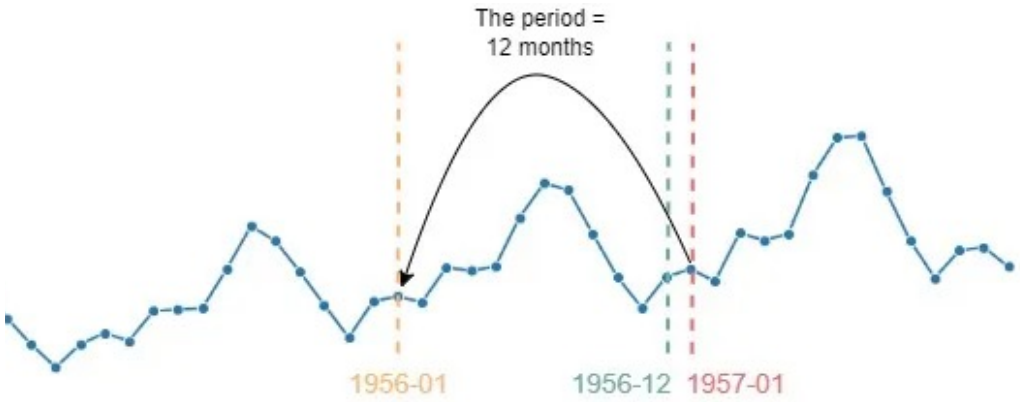
vs

Outputs: [$\overline{X}_{t+1}$, $\overline{X}_{t+2}$, $\overline{X}_{t+3}$, ...]

Encoder → Context Vectors → **Decoder**

[... , X_{t-2} , X_{t-1} , X_t]

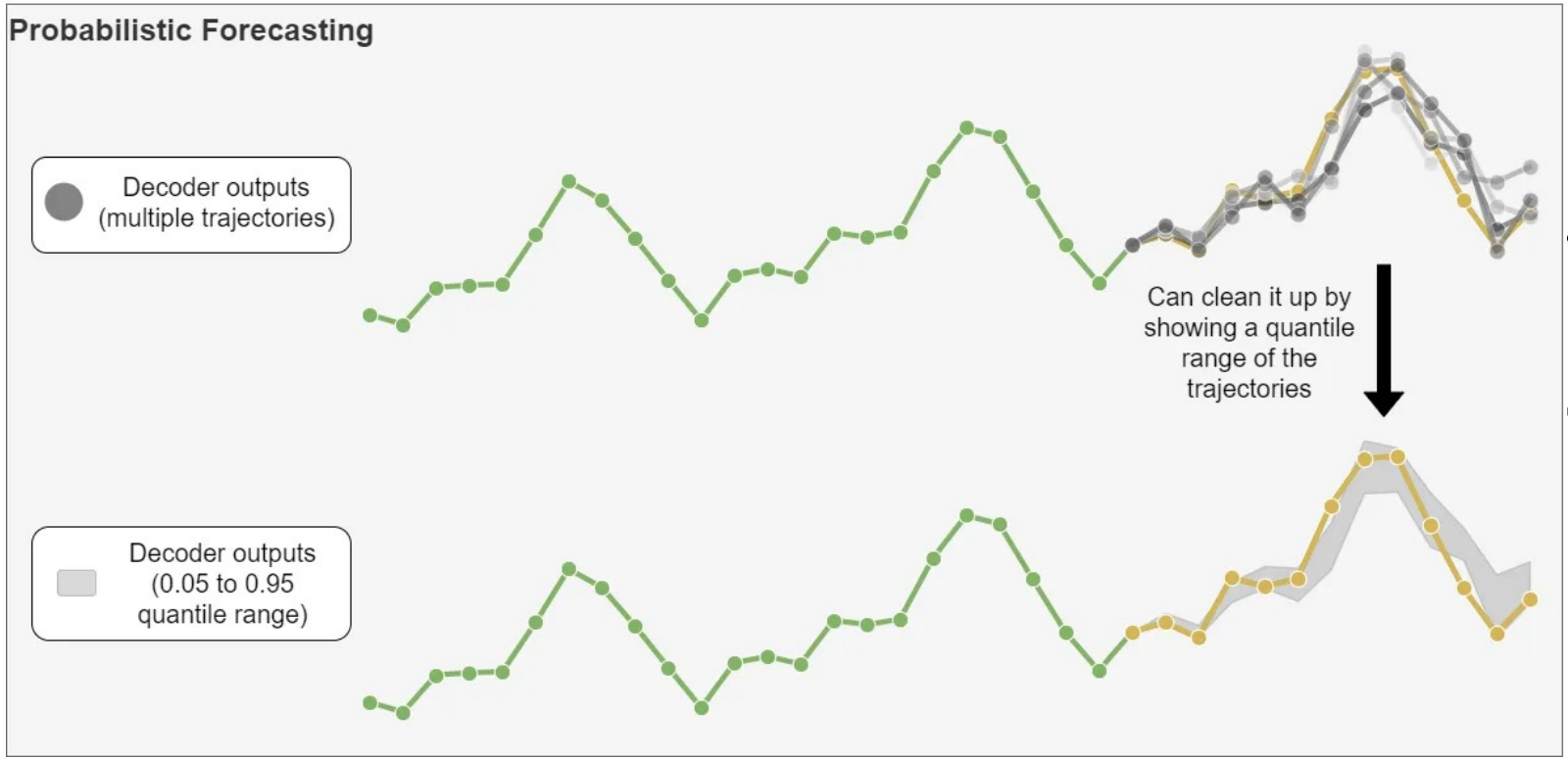


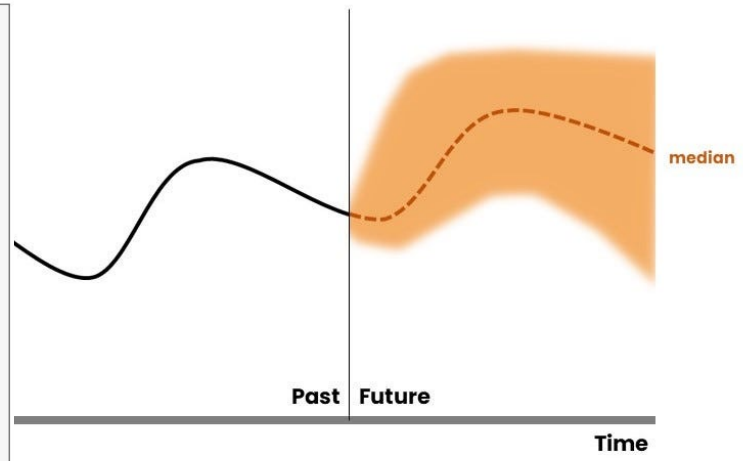


To assist in predicting the target value at 1957-01 from 1956-12, also take as input the target's corresponding previous period value at 1956-01

Prediction Details

Deterministic forecast



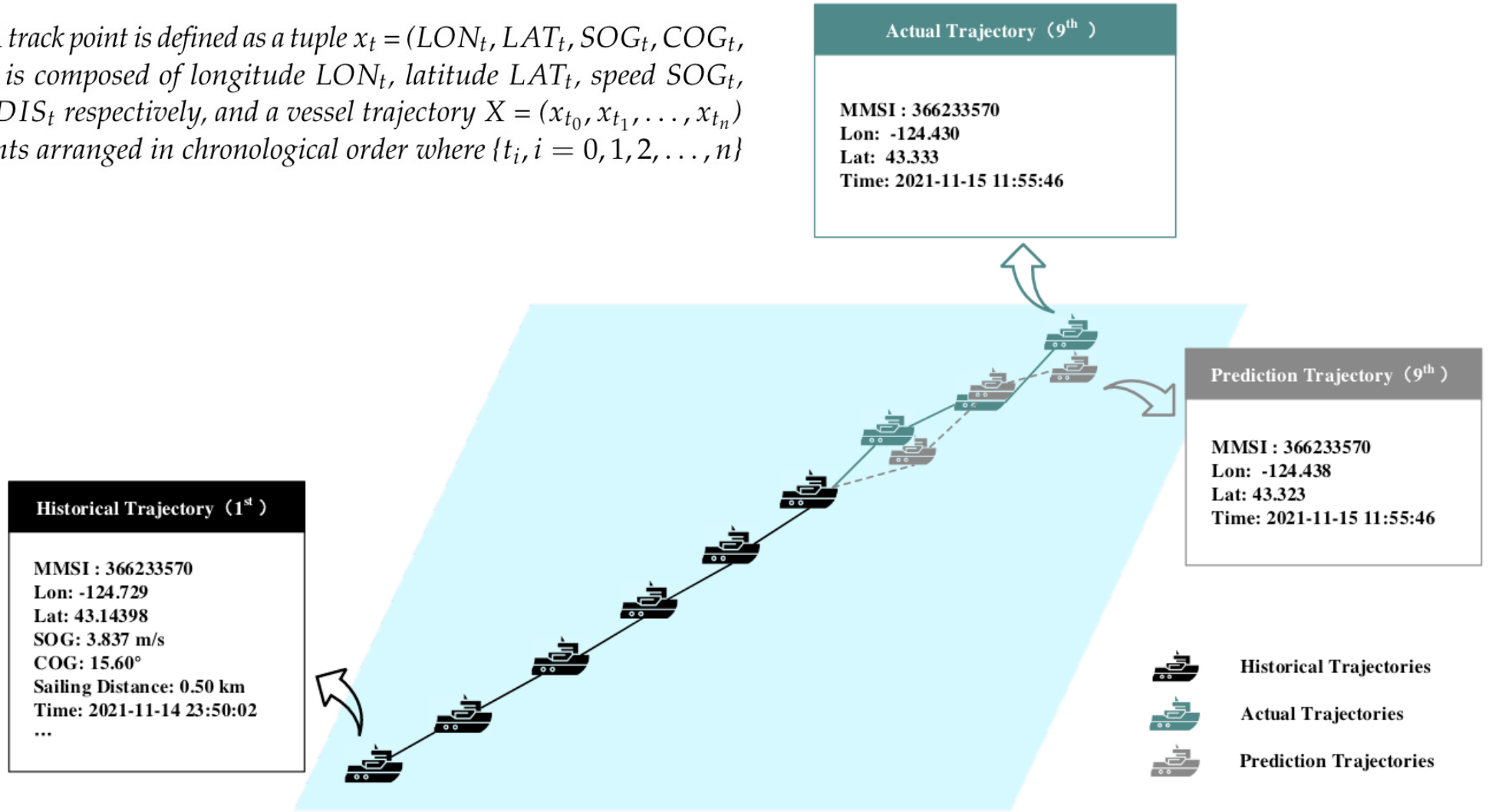


04

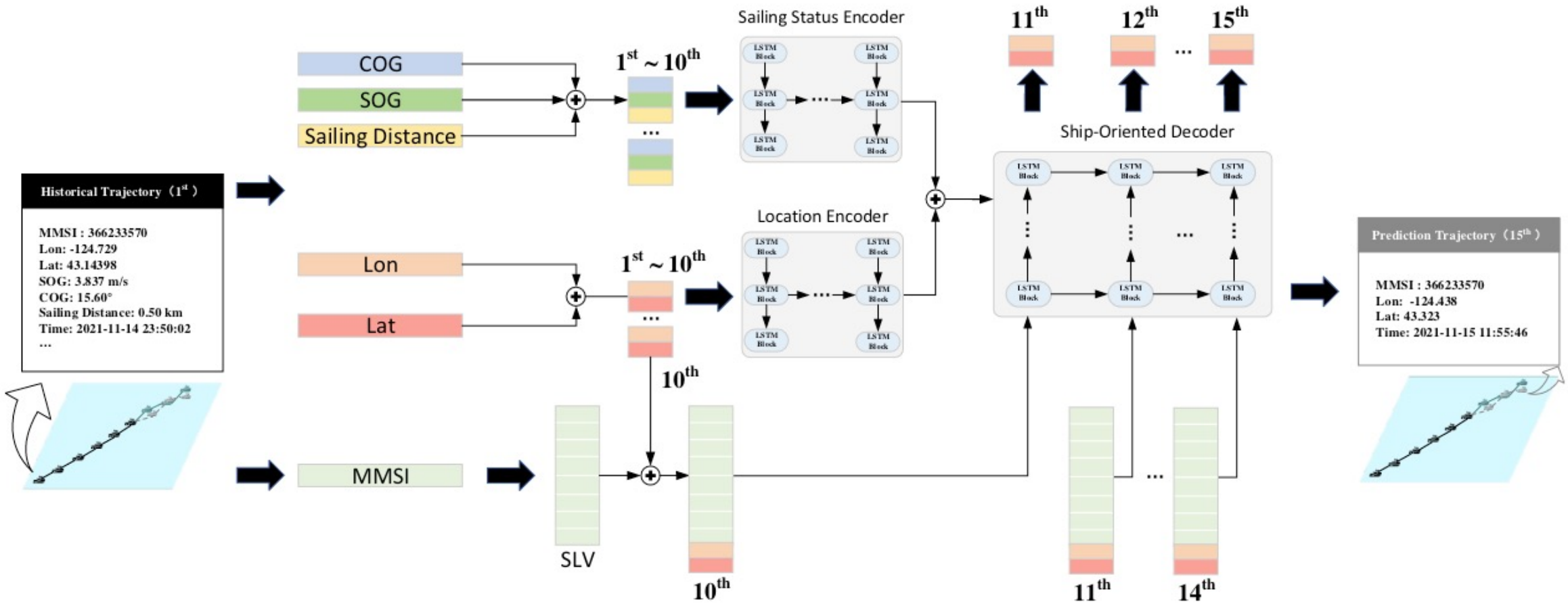
Vessel Trajectory Prediction

Background

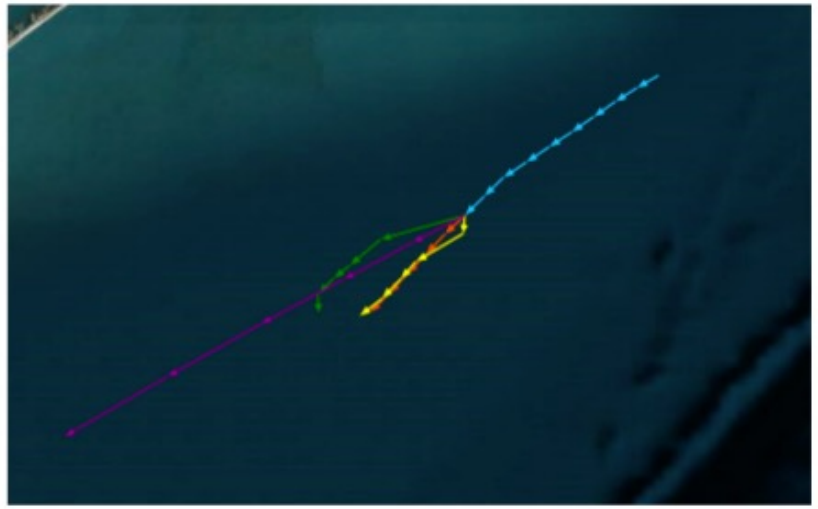
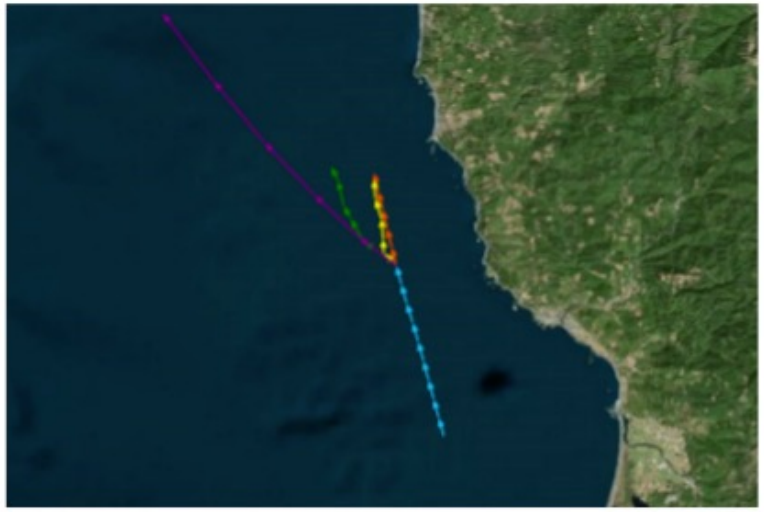
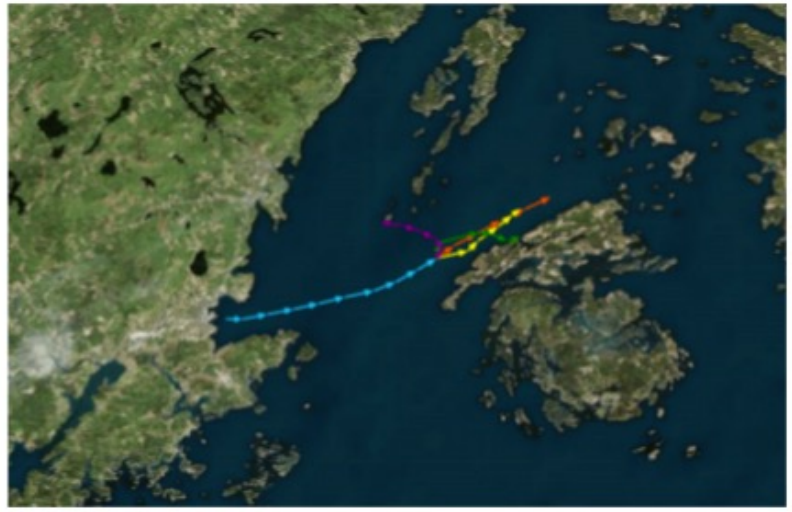
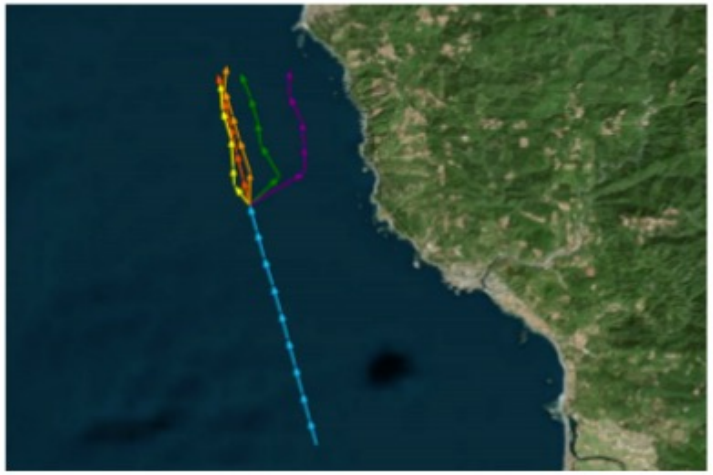
Definition 1. *(Vessel trajectory): A track point is defined as a tuple* $x_t = (LON_t, LAT_t, SOG_t, COG_t, DIS_t)$ *at the time of* t *in which* x_t *is composed of longitude* LON_t*, latitude* LAT_t*, speed* SOG_t*, course* COG_t*, and sailing distance* DIS_t *respectively, and a vessel trajectory* $X = (x_{t_0}, x_{t_1}, \ldots, x_{t_n})$ *is defined as a sequence of these points arranged in chronological order where* $\{t_i, i = 0, 1, 2, \ldots, n\}$ *is a set of timestamps.*



Double Encoder Seq2Seq Model



Application Scenarios

暨南大學
JINAN UNIVERSITY

Thanks